Applied Energy 306 (2022) 118048

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Building retrofit hurdle rates and risk aversion in energy efficiency investments

Yuan Lai [a], Sokratis Papadopoulos [b], Franz Fuerst [c], Gary Pivo [d], Jacob Sagi [e], Constantine E. Kontokosta [b,f,*]

[a] Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA
[b] Center for Urban Science & Progress, New York University, New York, NY, USA
[c] Department of Land Economy, University of Cambridge, Cambridge, UK
[d] School of Landscape Architecture and Planning, University of Arizona, Tucson, AZ, USA
[e] Kenan-Flagler Business School, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[f] Marron Institute of Urban Management, New York University, New York, NY, USA

ARTICLE INFO

Keywords:
Building energy
Retrofit investment
Energy efficiency
Financial analysis
Machine learning

ABSTRACT

Despite extensive empirical evidence of the environmental benefits of green buildings and the increasing urgency to reduce carbon emissions in cities, there has been limited widespread adoption of energy retrofit investments in existing buildings. In this paper, we empirically model financial returns to energy retrofit investments for more than 3600 multifamily and commercial buildings in New York City, using a comprehensive database of energy audits and renovation work extracted from city records using a natural language processing algorithm. Based on auditor cost and savings estimates, the median internal rate of return for adopted energy conservation measures is 21% for multifamily buildings and 25% for office properties. Logistic regression modeling demonstrates adoption rates are higher for office buildings than multifamily, and in both cases adopter buildings tend to be larger, higher value, and less energy efficient prior to retrofit implementation. The economically significant magnitudes of returns to adopted energy conservation measures raise important questions about why many property owners choose *not* to adopt. As such, we discuss incentive and regulatory mechanisms that can overcome financial and informational barriers to the adoption of energy efficiency measures.

1. Introduction

Retrofitting existing buildings has the potential to significantly reduce global energy use and carbon emissions, particularly in dense urban areas [1,2]. A broad range of international cities, including Tokyo, Singapore, Melbourne, London, and Toronto, have implemented policies designed to encourage or mandate more energy-efficient buildings [3]. Despite the positive impacts of reduced emissions and energy consumption, the pace of adoption of energy efficient practices and technologies has been slow, and substantial barriers – perceived and actual – persist [4–6]. These barriers, often considered to contribute to an energy efficiency gap [7], include both market failures and behavioral factors, such as information asymmetries between stakeholders, uncertainty over future savings, lack of knowledge about energy technologies, first-cost capital constraints, economic dis-incentives including the "split-incentive" problem, and fluctuating fuel pricing signals [8–10]. Mandatory energy disclosure and audit policies can overcome some of these challenges, and their recent proliferation across metropolitan areas has generated significant data on energy use and retrofit opportunities in buildings, led by New York City's (NYC) Local Laws 84 (LL84) and 87 (LL87) [11,12].

New energy disclosure, audit, and retro-commissioning requirements create detailed inventories of energy use profiles, building systems, and potential energy conservation measures (ECMs) [13]. NYC LL87 is the first city-wide building energy audit mandate for large office and multifamily buildings in the U.S. [14]. While several studies find that energy disclosure results in energy use reductions [15,16], mandatory audit policies have been found to have only a modest negative effect on building energy use over time [17]. Understanding the financial implications of the decision to adopt energy conservation technologies is a critical component of the broader push to increase energy efficiency in buildings. As such, NYC's mandatory audit policy provides a unique policy context to study the return on investment, or hurdle rate, that must be exceeded before retrofit investment is

* Correspondence to: Marron Institute of Urban Management and Center for Urban Science & Progress, USA.
E-mail address: ckontokosta@nyu.edu (C.E. Kontokosta).

https://doi.org/10.1016/j.apenergy.2021.118048
Received 21 June 2021; Received in revised form 24 September 2021; Accepted 9 October 2021

deemed profitable in the private sector. This knowledge could help to close the gap between perceived and actual financial risk associated with energy retrofits and subsequently guide education, regulations, or subsidy policies designed to promote energy use and carbon reductions. To address broader concerns about climate change, cities are beginning to introduce mandatory carbon reduction and energy efficiency targets for buildings [14,18]. Given regulatory and market pressures to improve energy efficiency, including new carbon emissions and efficiency targets,[1] building owners, investors, and policymakers need to understand the financial returns of the various pathways to energy use reductions through building retrofits in order to ensure incentives and penalties are sufficient to overcome existing barriers to large-scale retrofit adoption.

This paper examines a critical question about the link between building energy retrofit adoption and financial performance of energy efficiency investments using a unique, large-scale database of over 3600 office and multifamily buildings in New York City. We present a computational analysis of the potential investment return profiles for building retrofits across a range of building types and characteristics, and the associated likelihood of retrofit adoption using a logistic regression model. Data are collected and integrated from multiple sources, and include detailed information on energy use, building systems, financial metrics, construction permit records, and actual energy audit reports. We then calculate internal rates of return (IRRs) and develop net present value (NPV) curves for energy retrofit investments using reported audit data and permitted renovation work extracted using a natural language processing algorithm. The objectives of this study are to: (1) create a large-scale data repository of energy audit recommendations, building energy performance, building attributes, and renovation work using a natural language processing algorithm and data integration methods, (2) model and analyze the return on investment for various energy retrofit scenarios, including energy conservation measures adopted and those not adopted, (3) evaluate the financial drivers of the retrofit adoption decision, controlling for other factors that may influence the implementation of energy efficiency improvements, and (4) discuss applications of our analysis to advance energy efficiency and carbon reductions in global cities. Modeling the IRR and NPV for energy retrofit investments across heterogeneous property types and building characteristics provides the foundation for a data-driven understanding of the frictions hindering retrofit adoption and a more informed discussion of incentive and regulatory mechanisms to overcome financial and informational barriers.

2. Literature review

2.1. Energy retrofit decision-making and modeling

The decision to adopt an energy retrofit or energy efficiency technologies is driven by multiple factors [19,20]. These include behavioral attributes of key decision-makers in the organization (e.g. building owner, building management, shareholders) [21] and physical characteristics of the building itself, such as existing systems and technologies, building age, and building morphology [22]. Of particular significance are the economic and financial implications of an energy retrofit investment, which typically represent a primary constraint to energy retrofit adoption. These constraints include first, or upfront, capital costs of the ECM, the return on investment for individual or packages of ECMs, and the opportunity cost associated with retrofit investments as opposed to alternative investments. Of course, such considerations vary based on the time-dependent competitiveness of the local real estate market, the type and scale of the building, and the nature of the ownership entity [20,23]. Therefore, retrofit decision factors for single-family and low-density residential housing [24], for example, can differ significantly from those for large office or multifamily buildings in major urban cores. The incentives and barriers to retrofit adoption need to be understood in the context of building and ownership typologies and market segmentation [25]. Energy efficiency labeling – such as the U.S. Environmental Protection Agency's Energy Star certification and the U.S. Green Building Council's LEED rating – can potentially overcome some of these obstacles and has been shown to be associated with reduced energy consumption [19,26]. However, the effectiveness of such voluntary measures is limited as they only cover a subset of buildings, and selection bias tends to result in labeled buildings having above-average energy efficiency performance at the outset [27].

Researchers have developed a range of building energy retrofit decision support models to inform and optimize retrofit adoption. These can be grouped into simulation-based engineering models [28], data-driven and machine learning models [29], and hybrid approaches [30]. For instance, Chidiac et al. develop a screening approach that combines regression modeling of energy consumption with energy simulation to evaluate appropriate ECMs for Canadian office buildings [31]. Reflecting a more data-driven approach, Ali et al. use building performance data to estimate retrofit potentials across the residential building stock of Dublin [32]. Applying nine different machine learning algorithms, the authors identify key building characteristics, such as U-values of the envelope, influencing retrofit opportunities at scale. The role of financial constraints in retrofit decision-making are becoming an increasingly important consideration in model development. He et al. [33] develop an optimization algorithm to evaluate retrofit investment opportunities based present value and payback period financial metrics. The authors validate their model on a small sample of 27 buildings in the state of Delaware. Data limitations in previous energy audit decision studies can significantly limit the generalizability of the results and constrain opportunities to examine the potential financial implications for retrofits that were *not* adopted.

2.2. Financial returns to energy retrofits

Evidence has shown that energy efficient buildings are associated with higher rents and sales prices [34–36], occupancy rates [37], reduced operating costs, and, potentially, lower mortgage default risk [38,39]. These benefits contrast with the perceived under-allocation of resources for energy efficiency investments, resulting in what has been referred to as the energy efficiency gap [5,40]. Recent work on financial returns to building retrofits has primarily focused on macro-models of resource allocation for energy efficiency [41] or relied on small-sample cases studies [33] with limited diversity in building typology [42]. Similarly, theoretical optimization models are constrained by the lack of available data on actual ECMs adopted and the resultant return on those investments. Furthermore, data on specific retrofit opportunities *not* implemented are rarely available given the absence of widely-available audit databases [13]. Previous research has shown that the most significant barriers to retrofit adoption are information and market failures resulting in perceived or expected long payback periods on ECM investments and a lack of access to capital to fund implementation costs [11,20,40]. However, despite these theoretical and case study findings, there is little large-scale empirical understanding of the real-world potential return on investment of retrofit measures, how returns vary with individual ECMs and packages of ECMs across different building typologies, and the hurdle rate required by commercial building owners to invest in retrofits. This knowledge gap has nontrivial implications for the design, implementation, and evaluation of urban energy efficiency and climate policies.

[1] To address broader concerns about climate change, cities are beginning to introduce mandatory carbon reduction and energy efficiency targets for buildings. In NYC, the Climate Mobilization Act requires buildings over 2323 square meters (25,000 square feet) to reduce carbon emissions by 40% from 2005 levels by 2030 and 80% by 2050.

3. Data and methods

Fig. 1 summarizes our data integration and computational methodology. Using four years of energy audit reports provided by the NYC Mayor's Office of Sustainability and five years of construction permit records extracted from the NYC Department of Buildings (DOB), we first analyze a total of 22,230 ECM recommendations and their associated energy and cost savings estimates for 3632 individual office and multifamily buildings in New York City. We then conduct text mining to generate a dictionary of audit-recommended upgrades for each individual ECM category derived from the full audit report sample. To identify ECM adoption based on actual renovation activity subsequent to an audit, we match audit ECM recommendations with DOB building permit scope of work data for each of the 3632 buildings. Then, for each building, we estimate NPV and IRR for three scenarios representing return-maximizing, energy savings-maximizing, and balanced packages of ECMs. For buildings where audit recommendations were adopted, we calculate the IRR based on the bundle of adopted ECMs, and compare these values to the three potential adoption scenarios described above.

3.1. Energy retrofit investment net present value curves

Using the implementation (first) cost, energy savings, and annual cost savings data for individual ECM recommendations provided in each building's audit report, we compute the NPV for each ECM as follows:

$$NPV = \sum_{t=0}^{n} \frac{R_t}{(1+i)^t} \tag{1}$$

where n is the number of time periods of the investment, R_t is the net cash flow at period t, and i is the discount rate. For the purpose of this study, we assume $n = 15$ years and $i = 0.1$. The selection of the 15-year investment period is based on the average estimated useful lifespan of common ECM categories provided in the Advanced Energy Retrofit Guide by the Pacific Northwest National Laboratory and the U.S. Department of Energy [43]. The selection of discount rate is derived from PwC's Real Estate Investor Survey, 2nd Quarter 2018 data, which shows that rates ranged from 5.5% to 11.0% for 2015 (the median year of the data in our sample) for office buildings and from 5.0% to 10.0% for multifamily residential buildings. The average discount rate was 7.34% for office buildings and 7.24% for multifamily buildings. We select a 10% discount rate to reflect the higher risk premium associated with energy retrofit investments, while remaining within the survey range presented above. However, the selection of investment time horizon (n) and discount rate (i) can have significant implications for estimates of financial returns; therefore, we conduct a sensitivity analysis using 10-years, 15-years, and 20-years for n and 5.0%, 7.5% and 10.0% for i.

After calculating the NPV for individual ECMs, we are able to compute the cumulative NPV for all ECM recommendations for each building and plot the calculated values by cumulative energy savings. Fig. 2 shows the NPV/energy savings curve for a sample building, with each point indicated on the curve associated with a specific ECM. Note that we normalize both NPV and energy savings by building floor area to allow for comparison across building size. The order of the ECMs along the curve (from left to right) is based on the individual ECM's NPV, with the highest NPV first, then second highest, and so on. The curve presented in Fig. 2 is one of three commonly-identified retrofit investment NPV profiles, with the other two being a linear positive slope and a linear negative slope.

In this particular example, we see that the cumulative NPV curve peaks after two ECMs (specifically, HVAC controls and occupancy sensors for the lighting system), and the remaining ECMs are NPV negative. However, only the last recommendation (for conveying systems) causes the building's cumulative energy retrofit NPV to drop below zero. Calculating cumulative NPV/energy savings curves for each building in the dataset allows us to study inflection points in the curves, draw a more nuanced picture of the proposed ECMs' economic feasibility, focus on certain subsets of ECMs, and compute additional financial metrics. Based on these curves, we define three retrofit scenarios representing packages of recommended ECMs: NPV_{max}: the set of ECMs that maximize NPV, $NPV_{neutral}$: the set of ECMs yielding cumulative NPV close to or at zero,[1] and $EnergySavings_{max}$: all ECMs that would result in the greatest possible energy savings. For each scenario, we calculate the IRR for the identified bundle of ECMs. Moreover, based on the building's physical and energy use characteristics (age, gross floor area, energy use intensity, etc.), we further subset the data and study the aforementioned metrics by building sub-categories.

3.2. Text mining and audit-to-permit matching

Natural Language Processing (NLP) is a machine learning method for analyzing large collections of human-interpretable text data [44]. Computationally, NLP generates statistical measures by parsing, searching, counting, and summarizing frequency distributions of words, and further gains semantic insights such as frequently-mentioned words or topics. In this study, we analyze building permit descriptions submitted to the NYC Department of Buildings using the Natural Language Toolkit (NLTK), a widely-adopted NLP package in the Python coding environment [45].

Fig. 6 illustrates our computational workflow using NLP to detect and classify ECM implementation through audit and building permit matching. In the LL87 audit data, each ECM recommendation has a category-suggestion data structure. Each suggestion's description contains one or more human-readable sentences ("natural language"). We first group all ECM descriptions by ECM category to process relevant text. In step 1, we use part-of-speech (POS) tagging to clean the raw ECM descriptions by dropping conjunctions, determiners, pronouns, and punctuation. For each word, we calculate its frequency based on its total appearance as a function of all words in the description. Using these outputs, step 2 then generates ECM category-specific dictionaries by extracting text from auditors' recommendations (e.g., upgrade lighting to LED). Therefore, the final dictionary contains all unique words and their frequency associated with individual ECM recommendations.

For DOB building permit descriptions (step 3), we clean input text for the scope of work description using a similar process as step 1. Step 4 retrieves ECM recommendations extracted from audit reports performed prior to the permit application. We identify buildings in the sample with permitted alteration work subsequent to the date the audit was performed, based on the filing dates of the audit report and any construction permits in the DOB database. If a building has no post-audit permit record, we assume no renovation activity occurred in the building and thus no audit recommendations were adopted. It is possible, however, that the implementation of a particular ECM would not require the filing of a building permit; we discuss this scenario in more detail below. Step 5 uses the ECM dictionaries generated (output of step 2) from the audit reports to estimate the adoption likelihood for each ECM recommendation, according to its identified post-audit building permit description. Specifically, it compares the content between a permit description and a specific ECM recommendation using a word-matching algorithm that proceeds as follows: First, according to

[1] The $NPV_{neutral}$ scenario includes all ECMs that yield a cumulative NPV close to zero, such that the next recommended ECM (ranked by NPV) would make the cumulative NPV less than zero. Given this stepwise approach to including ECMs, the $NPV_{neutral}$ scenario often has a cumulative NPV greater than zero, resulting in IRRs higher than the discount rate.

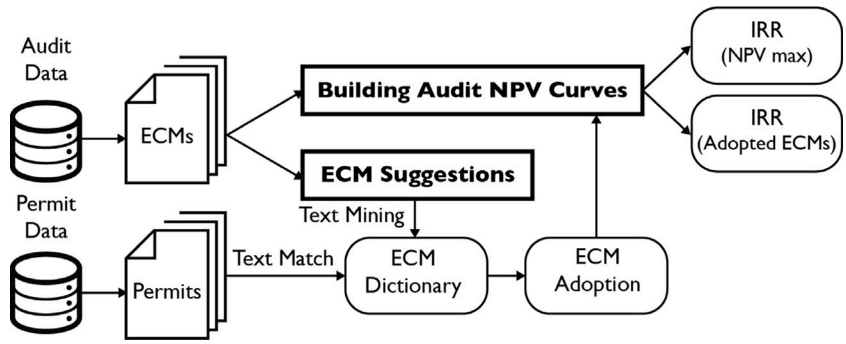

Fig. 1. Data processing and methodology.

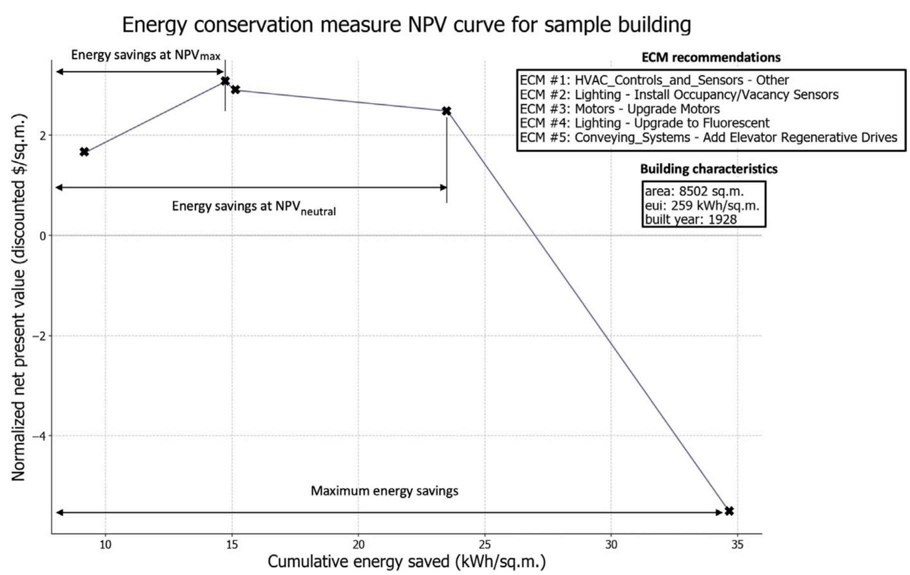

Fig. 2. Sample building NPV curve. ECM order is based on each ECM's NPV, ranked from highest to lowest (from left to right).

the ECM category, it associates the scope of work description with the ECM dictionary. Then, based on this dictionary, it identifies relevant words that appear in the permit description. Finally, it returns two new variables: (1) the total number of matched words and (2) a list of matched words. This approach quantifies the relationship between post-audit building permit descriptions and each ECM recommendation category from the audit report as an estimate of the likelihood of ECM adoption.[2]

3.3. Logistic regression model of retrofit adoption

We investigate the effects of relevant building and financial characteristics on the likelihood of adopting recommended ECMs using a multivariate logistic regression model. The dependent variable is a binary classification of retrofit adoption for each building in the sample (Y), equal to 1 for adoption and 0 for non-adoption. The independent variables include building typology (office vs. multifamily), building age, built area, building site EUI, first-cost of recommended ECMs, property value, estimated IRR (NPV max scenario), and potential energy savings (NPV max scenario). The mathematical expression of the logistic regression is:

$$Logit(P) = ln[P/(1-P)] = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (2)$$

where P is the probability of $Y = 1$ based on the set of independent variables X. Thus, the odds of the binary output $Y = 1$ based on the set of attributes X can be expressed as P/(1-P) and the $ln(P)$ is the natural log of the odds ratio (OR).

4. Results

Across the 3632 audit reports in our sample, we find the top five most commonly recommended ECM categories to be lighting (28%), domestic hot water (17%), envelope (13%), HVAC controls and sensors (10%), and distribution systems (8%) (Table 4). Fig. 3 presents a boxplot of the calculated simple payback period by ECM category based on auditors' estimates. The distribution of payback periods within each category are a result, in part, of the range of specific recommendations contained within each of the higher-order ECM categories (e.g., cooling system, conveying system) and the variance in auditor estimates.

4.1. Drivers of retrofit adoption

After matching LL87 audit data and DOB building permits by building (using the "BBL" building unique identifier), we find 1385 buildings

[2] For additional details on this methodology, please see Lai and Kontokosta [46].

Table 1
Comparison of audits, ECM recommendations, and building characteristics for non-adopters and adopters.

Building Type	Office		Multifamily	
Total Audits	405		3209	
Total ECM Suggestions	1988		17786	
	Non-Adopters	**Adopters**	**Non-Adopters**	**Adopters**
Number of Audits	277 (68%)	128 (32%)	2588 (81%) Condo = 314, Co-op = 2274	621 (19%) Condo = 95, Co-op = 526
Number of ECMs	1281 (64%)	707 (36%)	14248 (80%)	3538 (20%)
Median Built Year	1927	1928	1941	1942
Median Building Area (m^2)*	14602	16843	7515	9626
Median Site EUI (kWh/m^2)*	249	268	256	256
Median Value ($\$/m^2$)*	1184	1292	355	506
Median Energy Savings NPV max (kWh/m^2)*	11	13	21	17

NOTE: This table report results based on '75th perc +' scenario.
*Two-sample T-test significant at 95% level ($p \leq 0.05$).

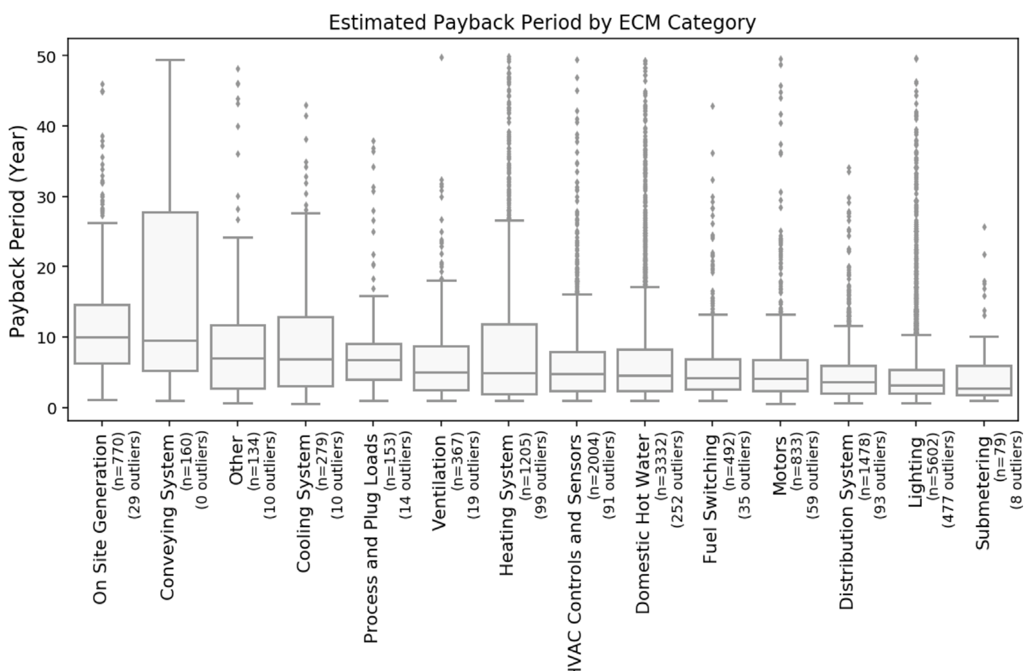

Fig. 3. Box-plot of estimated payback period distributions by ECM category sorted by average payback period.

with an audit *and* at least one building permit filed after the date the audit was conducted. We define this as a *post-audit alteration*. There are a total of 6,111 post-audit alterations since one building may file multiple alteration applications. For buildings with post-audit alterations, a total of 6,545 ECMs are matched between the audit reports and DOB permit descriptions, including lighting (n=2,028), domestic hot water (n=934), envelope (n=856), HVAC controls and sensors (n=634), distribution system (n=537), heating system (n=427), motors (n=234), fuel switching (n=175), cooling system (n=168), ventilation(n=160), on-site generation (n=147), conveying systems (n=77), process and plug loads (n=40), and sub-metering (n=33). For each ECM suggestion, our NLP algorithm retrieves associated building permit descriptions and identifies matched words based on the generated ECM-category dictionary. Using the distribution of total matched words, we define three different matching criteria. We use a 90th percentile threshold (matched words >= 6) as a conservative matching scenario (labeled as *90th perc*) and 75th percentile (matched words >= 3) as our base matching scenario (labeled as *75th perc*). Furthermore, most building permit descriptions do not report lighting improvements (e.g., upgrade bulbs to LED, install timers) since these actions may not involve work defined by the DOB as requiring a permit. Therefore, using the 75th percentile matching results, we define a third scenario by assuming the building also implemented recommended lighting ECMs that would not require a permit (labeled as *75th perc+*). According to the audit records, the "envelope" ECM category includes specific recommendations such as increasing roof insulation (23.0%), sealing doors (19.0%), replacing windows (15.6%), increasing wall insulation (13.4%), adding window films (8.3%) and sealing room AC (5.6%). Across all scenarios, we find over-matching for the "envelope" ECM category given the wide range of generic terms used to describe this category in the DOB permit scope of work descriptions (e.g., vocabularies include "floor", "wall", "door", and "window"). Therefore, given the limitations created by the lack of detail in work descriptions provided in the DOB database, we exclude this category from the IRR and NPV calculations.

We consider a building to be an energy retrofit "adopter" if there is at least one ECM match between the audit recommendations and post-audit DOB permit scope of work description. We compare the

Table 2
Logistic regression model results showing the likelihood of energy retrofit adoption based on building characteristics and audit recommendations.

	OR	Coef.	Std. Err.	z	P> |z|	[0.025	0.975]
Office	1.011	0.011	0.196	0.057	0.954	−0.373	0.396
Building Age	1.009	0.009	0.002	5.585	0.000	0.006	0.012
Building Area	1.000	0.000	0.000	0.537	0.591	−0.000	0.000
Site EUI	0.081	0.000	0.000	0.703	0.482	−0.000	0.000
Cost	1.002	−2.508	0.266	−9.433	0.000	−3.029	−1.987
Value	1.000	0.002	0.001	2.372	0.018	0.000	0.004
IRR[a]	1.127	0.119	0.244	0.489	0.625	−0.359	0.598
Energy Savings[a]	0.999	−0.000	0.000	−0.148	0.883	−0.000	0.000
Dependent Variable: adopted (1/0)					LLR p-value: 4.0118e−42		
Pseudo R-squared: 0.134					Classifier Accuracy: 0.58		

[a]Estimations are based on NPV max scenario.

number of audits and total number of ECMs recommended for office and multifamily buildings grouped by non-adopters and adopters based on the *75th perc+* matching criteria (Table 1). We also compare building characteristics, including built year, residential property ownership (condominium vs. co-operative), built area, and assessed value (in US$ per square meter) by merging with NYC Primary Land Use Tax Lot Output (PLUTO) data. Overall, office buildings have a higher adoption rate (32%) than multifamily buildings (19%). For multifamily buildings, results show the adoption rate in co-operatives (18%) is lower than condominiums (23%), possibly due to additional board approval requirement for building improvements in co-operative properties and the underlying financing structure of this ownership type. Two-sample t-tests indicate statistically significant differences in building area, initial energy use intensity (EUI), and median assessed value per square meter. For both office and multifamily, buildings that adopt ECM recommendations are found to be larger, higher value, and have higher potential energy savings (for office buildings only) than the ECMs identified in the NPV_{max} scenario. Although buildings that adopt tend to be newer, there is no statistically significant difference in built year.

According to our analysis of the LL84 energy performance database, energy use, measured as site EUI, is higher initially for adopter buildings than non-adopters in 2013, but decreases in the adopter buildings over the study period, as shown in Fig. 4. Between 2013 and 2017, EUI for adopter buildings decreased by approximately 3.5% for office and 1% for multifamily buildings. Non-adopter buildings, on the other hand, reported an increasing EUI, up by as much as 5.7% over the five-year time period (for additional analysis, please see Papadopoulos et al. [16]).

Results of the logistic regression model, shown in Table 3, indicate building age (p-value < 0.001, coefficient=0.009), ECM cost per square meter (p-value < 0.001, coefficient=−2.508), and property value per square meter (p-value < 0.05, coefficient=0.002) have statistically significant associations with the likelihood to adopt. Older buildings with higher property value are more likely to adopt ECM recommendations, holding other attributes constant. Furthermore, less-costly ECM recommendations are more likely to be adopted. Notably, estimated IRR and energy savings based on the audit recommendations are not statistically significant factors in the decision to adopt. This reinforces findings from the descriptive analysis of adopters and non-adopters, which indicates similar mean IRR values across the two groups. (See Table 4.)

4.2. Return on investment under multiple retrofit scenarios

Fig. 5 compares the calculated IRR distributions based on ECMs (1) included in the NPV_{max} scenario, (2) included in the $NPV_{neutral}$ scenario, and (3) those actually adopted using the *75th perc+* matching criterion. Median IRRs for the bundle of adopted ECMs are found to be 21% for multifamily and 25% for office. For both building types, the IRR of adopted ECMs has a lower mean and is negatively skewed relative to the $NPV_{neutral}$ scenario. In the discussion section, we elaborate on why this may be the case. From Table 2, we find that the median cost for adopted ECMs is $11.95 and $6.70 per square meter for office and multifamily, respectively, situating the first cost between the ECMs for the NPV_{max} and $NPV_{neutral}$ scenarios for multifamily and between the $NPV_{neutral}$ and maximum energy savings scenarios for office. For office buildings, the expected energy savings of adopted ECMs is 20.3 kWh per square meter, less than the expected savings from the $NPV_{neutral}$ scenario. (See Table 4.)

Multifamily buildings exhibit a similar pattern, with an estimated IRR for adopted ECMs of 21%, slightly below the IRR of the $NPV_{neutral}$ scenario of 22%. First costs of adopted ECMs are approximately $2.39 per square meter higher than the NPV_{max} scenario, but less than the $NPV_{neutral}$ ECM package. Expected energy savings are lower than those in the NPV_{max} and $NPV_{neutral}$ alternatives, indicating the adoption of improvements that may have higher first costs and lower energy savings over time, such as conveying (e.g. elevator) systems. For both building types, there is not a clear relationship between the variability in estimated payback period for a given ECM category (as a proxy for uncertainty in projected cost and savings) and its adoption. [3]

We also examine the extent to which additional ECMs, beyond those adopted, would have improved the expected return, referred to here as the "next-best" ECM. The next-best ECM is defined as the ECM with the highest NPV that was not implemented as part of the bundle of ECMs matched to the building's renovation permit scope of work. If the next-best ECM had been adopted, we find that the IRR would increase by 2% for multifamily properties, but would decline by 1% for office buildings. The most commonly identified next-best ECM for office buildings is fuel switching, a relatively high-cost investment that is dependent on infrastructure access to alternate fuel sources (e.g. natural gas) and on the price variability of different fuels. For multifamily buildings, distribution system improvements and fuel switching are found to be the among the next-best ECM alternatives based on NPV.

5. Discussion

Our results demonstrate a 21% median IRR for adopted retrofit investments for multifamily buildings and 25% for office buildings. Adopter buildings tend to be larger, higher value properties with higher initial EUI. Furthermore, we find that adopted ECM investments are associated with energy savings of approximately 7.6% and 5.9% for office and multifamily buildings, respectively. The adoption decision is driven by many factors, including capital constraints, behavioral influences, and uncertainty associated with energy and cost savings over time. The magnitude of the financial returns found here could potentially reflect a risk premium tied to the perceived uncertainty in the savings estimates provided in the audit reports and the opportunity cost of investing in retrofits over other, more traditional alternatives. Nonetheless, we find that returns to adopted ECMs are lower than what would be achievable for retrofit scenarios yielding the highest NPV. This is consistent with both market failure and behavioral explanations for the energy efficiency gap. First, it is possible that some adopter building owners emphasize energy savings, while sacrificing positive NPV investment options. For these owners, the ECM selection process may weight energy savings more heavily, even if the ECM does not increase the cumulative NPV of the investment because of higher first (implementation) cost. This is supported by the results of the relative energy savings associated with adopted ECMs when compared to the

[3] Fig. 7 presents the adoption rate for ECM categories plotted against the range in projected payback period, measured by the difference (in years) between the 5th percentile and 95th percentile payback period estimate for each ECM.

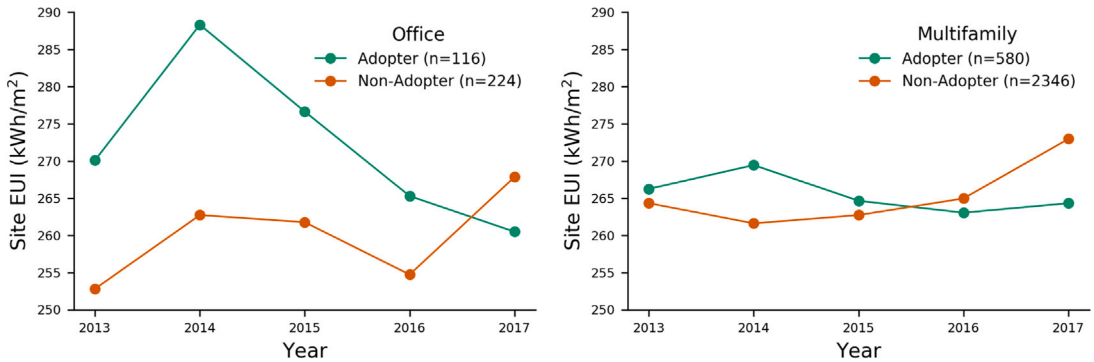

Fig. 4. Median energy performance (site EUI) over time (2013–2017) for office (left) and multifamily (right) buildings.

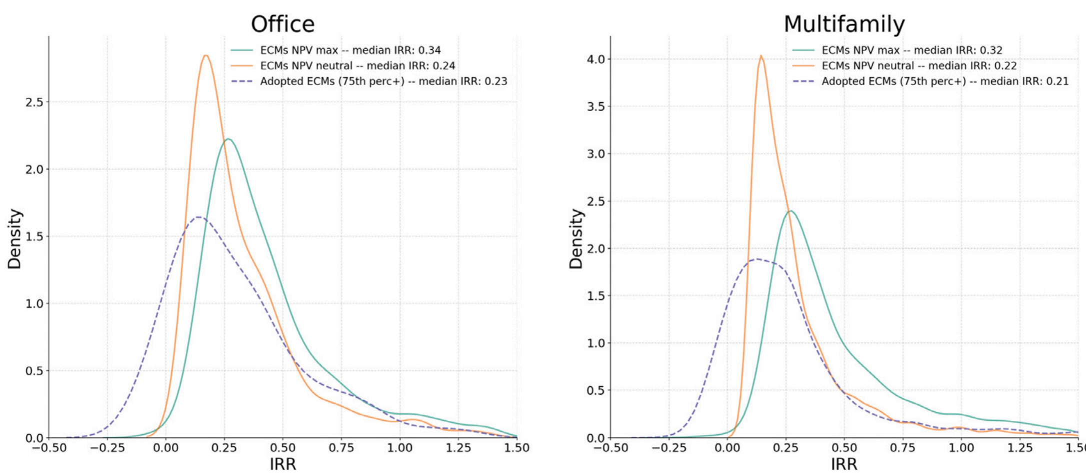

Fig. 5. Internal rate of return (IRR) for office and multifamily building energy retrofits, based on *NPVmax*, *NPVneutral*, and *adopted* scenarios.

Table 3
Comparative analysis of IRR, expected energy savings, and cost of ECM implementation, adopters and non-adopters.

IRR										
		NPV max scenario			NPV neutral scenario			Adopted ECMs		
		median	mean	std	median	mean	std	median	mean	std
Office	Non-adopter	0.31	0.38	0.20	0.24	0.29	0.18	–	–	–
	Adopter	0.31	0.38	0.19	0.24	0.30	0.18	0.25	0.28	0.22
Multifamily	Non-adopter	0.32	0.42	0.26	0.19	0.27	0.22	–	–	–
	Adopter	0.32	0.41	0.25	0.22	0.28	0.21	0.21	0.25	0.27
Energy Savings (kWh/m^2)										
		NPV max scenario			NPV neutral scenario			Adopted ECMs		
		median	mean	std	median	mean	std	median	mean	std
Office	Non-adopter	13.4	16.7	14.3	16.7	21.1	18.2	–	–	–
	Adopter	12	16.7	15.3	15.5	20.5	18.8	9.3	20.3	32.2
Multifamily	Non-adopter	23.3	27.1	20.9	31.1	37.2	28.9	–	–	–
	Adopter	15.6	22.2	20.3	22.4	30.2	28.9	10.0	15.1	20.0
Median First Cost ($/m^2)										
		NPV max scenario		NPV neutral scenario		Adopted ECMs		Max savings scenario		
Office	Non-adopter	6.46		8.50		–		16.47		
	Adopter	5.81		8.18		11.95		13.67		
Multifamily	Non-adopter	5.70		13.67		–		23.14		
	Adopter	4.31		8.50		6.70		16.04		

NPV_{max} scenarios. Second, if adopters' cost of capital is typically lower than 10%, then they will take up projects that appear to be negative NPV when discounted at 10%, thus biasing the adopter curve down.[4] Finally, behavioral factors, such as present bias or limited attention, certainly play a role in the energy retrofit adoption decision [47–49]. Many of these influences can be operationalized as economic considerations by, for instance, shifting discount rates to account for uncertainty in future savings. The analysis presented here provides the foundation for a deeper exploration into the relative significance of market failures and financial considerations as compared to behavioral barriers.

We also find that the "next-best" ECM would decrease the IRR of the aggregate retrofit investment by 1% for office, but increase the return by 2%, on average, for multifamily buildings. The next-best ECM for office buildings is determined to be fuel switching, which has high implementation costs and variable energy savings based on energy price fluctuations and the availability of alternate fuel source infrastructure. For multifamily properties, the next-best ECM is the distribution systems category, which can present challenges given constraints on access to individual apartments to do recommended work. The technical challenges and financial implications of the next-best ECM suggest that owners are balancing return and energy savings in the decision process.

To assess the accuracy of the algorithmic matching for ECM adoptions based on permit descriptions, we randomly select 30 samples five times (n = 150) to manually interpret the permit descriptions and then evaluate the accuracy for each matching result. The accuracy for each round is 0.90, 0.87, 0.90, 0.80, 0.83, resulting a mean accuracy of 0.86. Recall is estimated to be 0.92 and precision is 0.58. From the review of the full permit text description, we find a significant number of the observed false positives resulting from buildings that have filed many permits accumulating long text descriptions. This highlights a potential limitation in frequency-based matching techniques with varying text description lengths. Furthermore, approximately 20% of false positives were associated with the "envelope" ECM category, providing support for the exclusion of this category as described above.

We acknowledge that our approach has several limitations, primarily due to data sparsity and audit quality. A number of assumptions are made to estimate NPV of the various ECM scenarios, including discount rate and useful lifespan of the installed system or improvement. To add robustness to the analysis, we consider uncertainties based on distributions of input parameters using sensitivity analysis for discount rate and ECM lifespan. Different reporting systems (audit vs. permit) and data entry standards (auditor vs. contractor/architect/engineer) create uncertainties in text matching, which can lead to a mis-allocation of adopted ECMs. Building permit work descriptions are often vague and may not capture all ECM categories since several ECMs may not constitute work requiring a building permit. For example, a building owner often does not need to file a permit application for lighting improvements that involve bulb replacement/switching or minor repair work. This missing information may cause an underestimation of lighting ECM adoptions, although we account for this in our model through our matching thresholds. Data quality is also a significant concern for both energy audit reports and permit scope of work descriptions. We find inconsistent input formats, naming conventions, and misreported or erroneous savings and cost projections. A data standardization effort for energy audit reports is underway in NYC; however, this does not address the underlying issue of the reference data and metrics used by auditors to estimate future savings.

[4] Multifamily properties are typically underwritten using a lower discount rate than office properties. Correspondingly, it is worth noting that the disparity between the adopter and NPV_{max} or $NPV_{neutral}$ curves is greater for multifamily properties.

Although multiple agencies and organizations collect data related to building energy performance, energy audits, and renovation work, these efforts are largely siloed and constrained by sparse datasets representing single building types, regions, or portfolios. Our methodology can be used to better integrate audit data, building characteristics, and permit scope of work information into a unified energy performance database. To improve data reliability, consistency, and geographic coverage, we propose to develop a *National Retrofit Investment and Performance (NRIP)* database. This database would track building-level energy audits, implemented energy conservation measures and retrofit investments and their financial and energy performance metrics, and pre/post energy use profiles. The NRIP would integrate directly with the U.S. Department of Energy's Building Performance Database and other federal resources (such as EPA's Portfolio Manager), and provide a detailed repository for actual building audits and retrofit measures.

6. Conclusion

This study provides new insight into the return on investment for actual energy improvements put-in-place and a methodological foundation for a large-scale, nation-wide study of building energy retrofit adoption. In particular, we highlight three primary contributions of this study: First, individual in-depth case studies typically present as a mix of idiosyncratic characteristics and more general features. This means that any lessons derived from in-depth, small-sample or prototype building studies are not readily transferable to other buildings. Our study addresses this challenge by analyzing a large sample of approximately 3600 buildings derived from a mandatory energy audit policy. The sample consists of required audit reports and minimizes self-selection bias observed in previous studies of the audit-retrofit adoption relationship. Second, linking the energy audit database from NYC LL87 with a full record of construction and renovation activities from the NYC DOB provides us with a more complete and accurate picture of the retrofit process than simply observing the outcomes of the process post-retrofit. We also use novel computational methods to extract relevant data from building permits and match these to energy audit recommendations to capture those buildings that implemented audit recommendations, and those that did not. Finally, we estimate the rates of return for retrofit investments for both adopter and non-adopter buildings to determine the investment hurdle rates for retrofits. With this, we analyze the characteristics that make it more likely that a building owner will adopt a particular measure or bundle of measures. In the aggregate, we develop a better understanding of the financial implications of large-scale retrofit adoption. This information can be used in practice by policymakers for devising new incentives and regulatory mechanisms, while building owners can use it to support evidence-based investment decisions. Ultimately, our exploratory study is intended to contribute to the knowledge base that can address financial and informational barriers to energy efficiency in buildings.

As cities introduce more expansive regulations for building energy efficiency and carbon emissions reductions, a complete understanding of the financial implications of retrofit investments is needed to evaluate viable pathways toward near- and long-term sustainability goals. For building owners, our IRR and NPV models provide greater insight into the financial returns to individual ECMs and packages of ECMs. For policymakers, the analysis can be used to assess the economic feasibility of new and existing regulations, and determine where incentives can help overcome barriers to adoption. By identifying buildings that adopted energy efficiency investments, and quantifying the return on those investments, we are able to fill a critical gap in the understanding of energy efficiency retrofits in existing buildings.

CRediT authorship contribution statement

Yuan Lai: Software, Formal analysis, Data curation, Writing – original draft. **Sokratis Papadopoulos:** Software, Formal analysis, Data curation, Writing – review & editing. **Franz Fuerst:** Conceptualization, Methodology, Writing – review & editing. **Gary Pivo:** Conceptualization, Methodology, Writing – review & editing. **Jacob Sagi:** Conceptualization, Methodology, Writing – review & editing. **Constantine E. Kontokosta:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank the NYC Mayor's Office of Sustainability for sharing the Local Law 84 and Local Law 87 ata used in this study, and we appreciate the helpful feedback of Paul Mathew, Cindy Zhu, and Andrew White, among others. This material is based upon work suppored by the National Science Foundation under Grant No. 1653772. This research was supported by the Real Estate Research Institute (RERI) and Lawrence Berkeley National Laboratory (LBNL). All errors remain our own.

Appendix

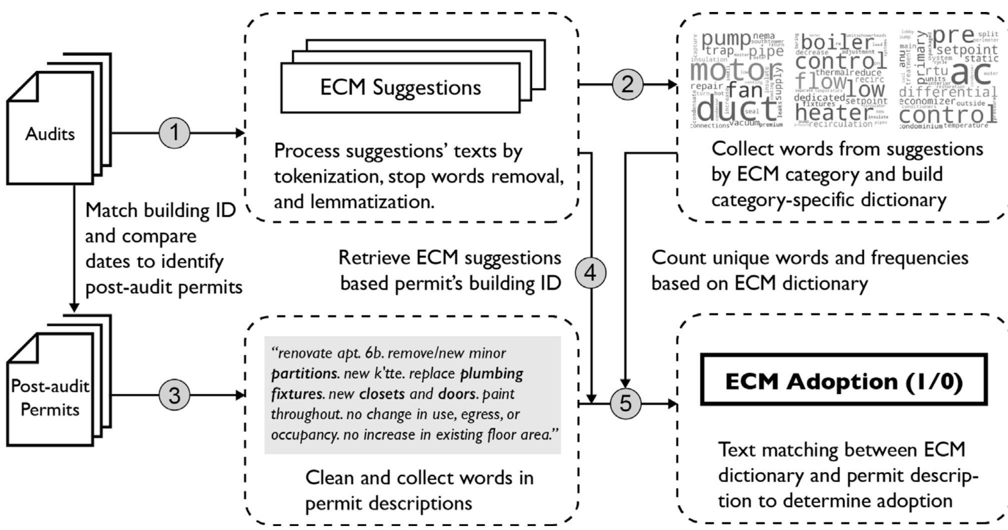

Fig. 6. Text data mining and natural language processing framework for ECM adoption classification.

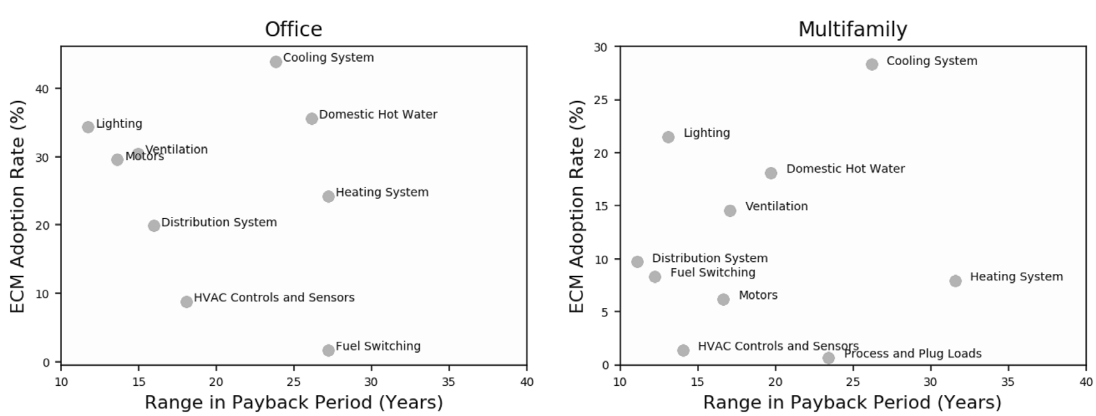

Fig. 7. Payback period range (from 5th and 95th percentiles) plotted against the rate of adoption for each ECM category.

Table 4
Summary of ECM categories and sub-recommendations.

ECM category	Suggestions (%)
On Site Generation (n = 770)	Install solar/photovoltaic (69.59%) Install co-generation plant (27.96%)
Conveying Systems (n = 160)	Add elevator regenerative drives (14.2%) Upgrade motors (31.5%) Upgrade controls (12.96%), Other (40.7%)
Cooling System (n = 279)	Replace package units (12.37%), Other (35.7%) Upgrade packaged units (10.70%), Upgrade chillers (8.36%) Add economizer cycle (8.36%) Add or upgrade cooling tower (6.35%)
Process and Plug Loads (n = 153)	Replace washing machines (38.99%), Other (52.83%) replace clothes dryers (3.14%) Automatic shutdown/sleep mode for computers (1.89%)
Ventilation (n = 367)	Other (32.25%), Install demand control ventilation (19.0%) Install exhaust fan timers (17.5%), Install CAR dampers (15.8%) Upgrade fan/ air handlers (7.0%), Upgrade exhaust fans (5.0%)
Heating System (n = 1205)	Upgrade burner (37.5%), Heating boiler upgrade (25.0%) Insulate vacuum pump assembly (12.5%) BMS/EMS installation (12.5%) Install indoor temperature sensors (12.5%)
HVAC Controls & Sensors (n = 2004)	Install or upgrade EMS/BMS (42.0%), Install TRVs (24.2%) Change Set Points/Setbacks — Heating (12.3%), Other (5.2%) Install indoor sensors (5.3%), Heat watch (3.7%) Install programmable thermostats (1.9%)
Domestic Hot Water (n = 3332)	Separate DHW from heating (32.8%) Install low-flow aerators (26.9%) Other (8.0%), Install low-flow showerheads (7.8%) Insulate DHW piping (6.8%), Install DHW controls (6.6%) Decrease DHW temperature (3.5%), Upgrade DHW boiler (2.2%) Low flow fixtures (1.3%)
Fuel Switching (n = 492)	#6 oil or #4 oil to natural gas (58.5%) #2 oil to natural gas (27.4%) #6 to dual fuel (4.8%) District steam to on-site generation (2.9%) Utility steam to on-site generation (2.5%)
Motors (n = 833)	Install VFDs (55.6%), Upgrade motors (35.9%) Other (4.7%), Remove motors (2.8%)
Distribution Systems (n = 1478)	Insulate pipes (80.0%), Other (20.0%)
Lighting (n = 5602)	Upgrade to LED (58.8%), Other (10.7%) Upgrade to fluorescent (6.6%) Install occupancy/vacancy sensors (6.6%) Upgrade exterior lighting (6.3%), Install bi-level lighting (1.7%) install photocell control (1.7%)
Submetering (n = 79)	LBS smart meters (50.4%) Install submetering (45.3%)

References

[1] Koomey JG, Martin NC, Brown M, Price LK, Levine MD. Costs of reducing carbon emissions: US building sector scenarios. Energy Policy 1998;26(5):433–40.

[2] Ma J, Qin J, Salsbury T, Xu P. Demand reduction in building energy systems based on economic model predictive control. Chem Eng Sci 2012;67(1):92–100.

[3] Trencher G, Broto VC, Takagi T, Sprigings Z, Nishida Y, Yarime M. Innovative policy practices to advance building energy efficiency and retrofitting: Approaches, impacts and challenges in ten C40 cities. Environ Sci Policy 2016;66:353–65.

[4] Liang J, Qiu Y, James T, Ruddell BL, Dalrymple M, Earl S, et al. Do energy retrofits work? Evidence from commercial and residential buildings in phoenix. J Environ Econ Manag 2018;92:726–43.

[5] Jaffe AB, Stavins RN. The energy-efficiency gap what does it mean? Energy Policy 1994;22(10):804–10.

[6] Ardente F, Beccali M, Cellura M, Mistretta M. Energy and environmental benefits in public buildings as a result of retrofit actions. Renew Sustain Energy Rev 2011;15(1):460–70.

[7] Allcott H, Greenstone M. Is there an energy efficiency gap? J Econ Perspect 2012;26(1):3–28.

[8] Mills E, Kromer S, Weiss G, Mathew PA. From volatility to value: analysing and managing financial and performance risk in energy savings projects. Energy Policy 2006;34(2):188–99.

[9] Palmer K, Walls M. Limited attention and the residential energy efficiency gap. Amer Econ Rev 2015;105(5):192–5.

[10] Ryghaug M, Sørensen KH. How energy efficiency fails in the building industry. Energy Policy 2009;37(3):984–91.

[11] Palmer K, Walls M. Using information to close the energy efficiency gap: a review of benchmarking and disclosure ordinances. Energy Effic 2017;10(3):673–91.

[12] Kontokosta CE. Energy disclosure, market behavior, and the building data ecosystem. Ann New York Acad Sci 2013;1295(1):34–43.

[13] Mathew PA, Dunn LN, Sohn MD, Mercado A, Custudio C, Walter T. Big-data for building energy performance: Lessons from assembling a very large national database of building energy use. Appl Energy 2015;140:85–93.

[14] Wyman KM, Spiegel-Feld D. The urban environmental renaissance. Calif Law Rev 2020;108:305.

[15] Meng T, Hsu D, Han A. Estimating energy savings from benchmarking policies in New York City. Energy 2017;133:415–23.

[16] Papadopoulos S, Bonczak B, Kontokosta CE. Pattern recognition in building energy performance over time using energy benchmarking data. Appl Energy 2018;221:576–86.

[17] Kontokosta CE, Spiegel-Feld D, Papadopoulos S. The impact of mandatory energy audits on building energy use. Nat. Energy 2020;5(4):309–16.

[18] Nadel S, Hinge A. Mandatory building performance standards: A key policy for achieving climate goals. Washington, DC, USA: American Council for An Energy Efficient Economy; 2020.

[19] Darko A, Chan APC, Ameyaw EE, He B-J, Olanipekun AO. Examining issues influencing green building technologies adoption: The United States green building experts' perspectives. Energy Build 2017;144:320–32.

[20] Kontokosta CE. Modeling the energy retrofit decision in commercial office buildings. Energy Build 2016;131:1–20.
[21] Christensen PH, Robinson SJ, Simons RA. The influence of energy considerations on decision making by institutional real estate owners in the US. Renew Sustain Energy Rev 2018;94:275–84.
[22] Li X, Ying Y, Xu X, Wang Y, Hussain SA, Hong T, Wang W. Identifying key determinants for building energy analysis from urban building datasets. Build Environ 2020;181:107114.
[23] Gliedt T, Hoicka CE. Energy upgrades as financial or strategic investment? Energy Star property owners and managers improving building energy performance. Appl Energy 2015;147:430–43.
[24] Kastner I, Stern PC. Examining the decision-making processes behind household energy investments: A review. Energy Res Soc Sci 2015;10:72–89.
[25] Hoicka CE, Parker P, Andrey J. Residential energy efficiency retrofits: How program design affects participation and outcomes. Energy Policy 2014;65:594–607.
[26] Asensio O, Delmas M. The effectiveness of US energy efficiency building labels. Nat Energy 2017;2(4):1–9.
[27] Walls M. Energy efficiency: Building labels lead to savings. Nat Energy 2017;2(4):1–2.
[28] Coakley D, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renew Sustain Energy Rev 2014;37:123–41.
[29] Deb C, Schlueter A. Review of data-driven energy modelling techniques for building retrofit. Renew Sustain Energy Rev 2021;144:110990.
[30] Hong T, Chen Y, Luo X, Luo N, Lee SH. Ten questions on urban building energy modeling. Build Environ 2020;168:106508.
[31] Chidiac S, Catania E, Morofsky E, Foo S. A screening methodology for implementing cost effective energy retrofit measures in Canadian office buildings. Energy Build 2011;43(2–3):614–20.
[32] Ali U, Shamsi MH, Bohacek M, Hoare C, Purcell K, Mangina E, et al. A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings. Appl Energy 2020;267:114861.
[33] He Y, Liao N, Bi J, Guo L. Investment decision-making optimization of energy efficiency retrofit measures in multiple buildings under financing budgetary restraint. J Cleaner Prod 2019;215:1078–94.
[34] Eichholtz P, Kok N, Quigley JM. Doing well by doing good? Green office buildings. Amer Econ Rev 2010;100(5):2492–509.
[35] Fuerst F, McAllister P. Green noise or green value? Measuring the effects of environmental certification on office values. Real Estate Econ 2011;39(1):45–69.
[36] Im J, Seo Y, Cetin KS, Singh J. Energy efficiency in US residential rental housing: Adoption rates and impact on rent. Appl Energy 2017;205:1021–33.
[37] Carlson K, Pressnail KD. Value impacts of energy efficiency retrofits on commercial office buildings in Toronto, Canada. Energy Build 2018;162:154–62.
[38] An X, Pivo G. Green buildings in commercial mortgage-backed securities: The effects of LEED and energy star certification on default risk and loan terms. Real Estate Econ 2020;48(1):7–42.
[39] Mathew P, Issler P, Wallace N. Should commercial mortgage lenders care about energy efficiency? lessons from a pilot study. Energy Policy 2021;150:112137.
[40] Gillingham K, Palmer K. Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence. Rev Environ Econ Policy 2014;8(1):18–38.
[41] Wang Z, Li H, Zhang B, Tian X, Zhao H, Bai Z. The greater the investment, the greater the loss?–Resource traps in Building energy efficiency retrofit (BEER) market. Resour Conserv Recy 2021;168:105459.
[42] Qu K, Chen X, Wang Y, Calautit J, Riffat S, Cui X. Comprehensive energy, economic and thermal comfort assessments for the passive energy retrofit of historical buildings-A case study of a late nineteenth-century Victorian house renovation in the UK. Energy 2021;220:119646.
[43] Liu G, Liu B, Wang W, Zhang J, Athalye RA, Moser D, et al. Advanced energy retrofit guide office buildings. Technical report, Richland, WA (United States: Pacific Northwest National Lab.(PNNL); 2011.
[44] Bird S, Klein E, Loper E. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. O'Reilly Media, Inc.; 2009.
[45] Project N. Natural language toolkit. 2021, URL https://www.nltk.org.
[46] Lai Y, Kontokosta CE. Topic modeling to discover the thematic structure and spatial-temporal patterns of building renovation and adaptive reuse in cities. Comput Environ Urban Syst 2019;78:101383.
[47] Schubert R, Stadelmann M. Energy-using durables–why consumers refrain from economically optimal choices. Front Energy Res 2015;3:7.
[48] Fuerst F, Singh R. How present bias forestalls energy efficiency upgrades: A study of household appliance purchases in India. J Cleaner Prod 2018;186:558–69.
[49] Schleich J, Gassmann X, Meissner T, Faure C. A large-scale test of the effects of time discounting, risk aversion, loss aversion, and present bias on household adoption of energy-efficient technologies. Energy Econ 2019;80:377–93.